IT Unit 3

Topic 2

# Ch. 3, Data Analytics: Drawing Conclusions, Part 2

**Identifying patterns and relationships between data**, 152-165

1. What is the purpose of using statistics when working with large data sets?

*Convert large quantities of raw data into small, informative, meaningful summaries. Allow patterns & relationships to be identified.*

1. List the 3 most useful statistical concepts from table 3.7, p 152.

* Average: the sum of data divided by the among of data pieces
* Standard Deviation: is how far away a value is away from the mean
* Correlation: do two dots show the same trend, as they might be connected

**Basic statistics**

1. Distinguish between the three methods of calculating an average.

* Mean: the total of a group of values divided by the among of value in that group
* Median : the middle value in a group of values
* Mode: the most frequent Value

1. What is the purpose of the standard deviation?

* How consistent a piece of data is compare to the mean of a data set?

**Correlation and causality**, p 156

1. Why do you need to be careful when looking at patterns in data in terms of cause and effect?

* People can miss interrupt graphs by thinking there is a causal relationship, trends occur when to sets of data are connected.

**Data visualisations, p 157**

1. What are data visualisations used for?

* To make interrupting data easier by using lines, shapes and colours to represent data

**Queries and searches,** p 159

1. How can queries, searches, filtering and sorting be used when analysing large data sets?

* Large sets of data may present a lot of information which may not be useful so using filters can help find interesting data that will be useful.

**Conditional formatting,** p 164

1. Explain the process of conditional formatting.

* Allows you to change the appearance of your data automatically based on its current value in both spreadsheet and database.
* Just changing values